THE EMBO JOURNAL

Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons

Jérôme Zervudacki, Agnès Yu, Delase Amesefe, Jingyu Wang, Jan Drouaud, Lionel Navarro and Angélique Deleris.

Review timeline:

Submission date:	24th October 2017
Editorial Decision:	17th November 2017
Revision received:	16th February 2018
Editorial Decision:	20th March 2018
Revision received:	13th April 2018
Accepted:	17th April 2018

Editor: Anne Nielsen

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 17th November 2017

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are shown below.

As you will see from the reports, our referees all express interest in the findings reported in your manuscript and support publication here, following adequate revision to address a number of major and minor points.

Given the referees' positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all three reviewers. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

For the revised manuscript I would particularly ask you to focus your efforts on the following points:
- strengthen the reproducibility and significance of the effects reported (refs #1 and #2). This can be done via additional replicas, more rigorous testing and by extended analysis of the already available data.
- extend the generality of the findings, both with respect to PRC2 targeting (ref #1 point 1) and the expression of immune-responsive genes (ref #3 points 1 and 2)
-> ref #3 also makes good suggestions to experiments that would test the broader functional/biological contribution from the SoloLTR in the plant response to pathogens (points 4 and 5) and I'd encourage you to extend the study in that direction.
- in addition, the refs all point to a number of minor points (references, methods, descriptions) that are ambiguous at the moment and need to be extended/clarified

---

REFEREE REPORTS.

Referee #1:

The authors of the manuscript "Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons" describe the involvement of the retrotransposon AtCOPIA93 (EVD) in the response to biotic stress. By using an LTR::GUS reporter line, the authors show that this element is transcriptionally regulated upon bacterial and elicitor treatment. They furthermore show that the regulation of the endogenous element is controlled by DNA and PRC2-mediated H3K27me3 histone methylation. Interestingly, the LTR of these elements harbors cis regulatory motives that regulate the transcriptional activity of neighboring genes, including the resistance gene RPP4, which the authors investigate in more detail. This manuscript is a nice example of the effects of retrotransposon insertions on neighboring gene expression upon biotic stress and provides evidence that LTRs harbor cis-regulatory motifs important for regulation. I have some critical comments aiming to improve this manuscript:

1. It would be interesting to know whether the LTR::GUS reporter is also targeted by H3K27me3. Since PRC2 targeting is largely determined by cis elements, one would expect the transgene to become targeted by H3K27me3. If not, this would argue against a cis-element-mediated PRC2 targeting of this LTR, which would be of interest in respect to currently unknown targeting mechanisms of TEs by PRC2.
2. The relevance of the double layer of regulation of EVD by DNA methylation and H3K27me3 has not been addressed. Investigating expression of EVD in clf met1 or clf ddm1 double mutants would be helpful to address this question.
3. It would strongly increase the value of the manuscript if the authors could test whether PAMP elicitation does induce new insertions of EVD. Since the authors have the material, doing a transposon display should not be that difficult to address this question.
4. Data shown in Fig. 1B, it is unclear how many lines were analyzed.
5. In figures 2B, 4A and 4B the variation between the experiments is quite high. A statistical treatment should be performed to test whether the observed differences are significant.
6. In figure 3B, the genes analyzed are not mentioned in the main text.
7. In Fig. 3A the authors show data of ChIP-Chip experiments. In the Supplement they should include data of published ChIP seq experiments (for wild-type), to add additional support for the presence of H3K27me3 on the LTR of EVD.
8. Fig. 3D: It unclear what exactly is plotted and what the points refer to.
9. I find it confusing that the authors use the name EVD and ATCOPIA93 synonymously. It would improve clarity to only use the name EVD as it has been previously published. It is standard in the field to use the gene name that has been published first; the same rule should be applied here for a TE.
10. References are not correct:
-page 2: "The maintenance of CHH methylation..." Deleris et al., do not address the mechanism of CHH methylation; Zemach et al., 2013 should be correctly cited here.
-page 6: Ordonnez et al., 2014 is not present in the reference list and also not the correct reference in this context.
11. Page 7, rephrase sentence :"...EVD almost exclusively is reactivated..."to "it was mostly EVD that became reactivated in the clf-elicited background."
12. Page 8, rephrase "---detection of transcription was detected..."

Referee #2:

The manuscript by Deleris et al., "Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons", describes the identification and functional characterization of a TE-derived pathogen-responsive promoter in Arabidopsis. Transgenic reporters reveal transcriptional activation of the LTR of a Copia element upon exposure to several pathogens, depending on lack of epigenetic silencing components and the presence of W box domains. The responsiveness seems to be determined by DNA methylation and Polycomb-based histone methylation. A soloLTR derived from the TE was found in the promoter of a pathogen-responsive gene.

These findings are interesting, and the work is largely well designed and performed. It provides another example for the potential role of TEs as regulatory elements, and it will be of interest to a broader readership. There are a few points that should be discussed.
1. Reproducibility between experiments is poor, as becomes evident from the lack of error bars in expression and ChIP experiments. However, the authors provide replicates as separate figures (e.g. 2A and EV2), but these often show substantial differences, in this case in kinetics as well as in the degree of response. Is this classical for pathogen infection experiments?
2. To judge the role of DNA methylation at the EVD LTR, the number of Cs in the different contexts should be shown. The information can be extracted from Figure EV1A but a close look reveals that (1) the first 2 of 7 CG positions in Figure 1 A are not within the region marked as LTR (and the Discussion mentions 5 sites, there are also only 6 CHG sites), and (2) less than half of all CHH sites are methylated at all, and most in just 1 of the 17 clones analysed. Other Copia LTRs are also characterized by low CG/CHG frequencies, and this special feature should be better discussed.
3. The analysis of just one transgenic line by bisulfite sequencing is unsatisfying. The added semi-quantitative analysis at one CG position by restriction digest analysis in Figure EV1B indicates no substantial methylation at the transgene also in other lines, but these data are weak.
4. H3K27me3 analysis: why was the transgenic line not included? As it was not DNA methylated (at least in the line analysed), it would have been informative to ask whether the LTR attracts the histone methylation independent of DNA methylation. If so, it could have revealed whether and how this modification eventually changes upon infiltration of flg22 or Ptodelta28E. This could have complemented the comparison between wt and clf, especially as clf leaves are morphologically quite different, to begin with, whereas the transgenic line has a normal leaf phenotype.
5. Role of the soloLTR5: it would have been much more elegant to eliminate the LTR from the RPP4 gene with gene editing, to avoid position effects in different transgenic lines. However, I accept the data from pools of primary transformants with ectopic insertions of the two different constructs in the rpp4 mutant background and their significant difference (although the different pool sizes in Fig. 5C, 33 versus 18, weaken the statistics).
6. The analysis of the role of the W box motifs in the LTR is inconsistent: while deletion of W2 in the GUS reporter seems to reduce the responsiveness more than that of W1 (Fig. 1E), only W1 was deleted in the soloLTR of the RPP4 construct (Fig. 5 B and D). Why?
7. P. 4 line 8 from bottom: "instead, we generated.." is misleading, make it clear that you just re-defined the purpose of the reporter line described in the first sentence of the paragraph.
8. Coexistence and cooperation of H3K27 methylation and DNA methylation in repression is not as novel as claimed here. A lot of literature is available in cancer epigenetics and should at least be mentioned.
9. More care could be taken to cite the correct primary references. Just one example: impaired CG methylation in met1 mutants was described prior to Ordonez et al. 2013.
10. Provide a reference for the Tanaka method in Material and Methods.
11. Editing of the manuscript by a native English speaker could improve grammar and punctuation.

Referee #3:

This analysis provides evidence of a causal role of transposons in regulation of plant immunity gene expression. Previous studies have shown that TN mobility is a driver of genome evolution and that TNs can be mobilized in response to various stresses. There is also evidence that TN insertions in the genome attract specific epigenetic marks to limit their activity and these marks can spread to cis-regulation of other genes. Here, the authors examine the effect of the epigenetic environment - DNA methylation and associated H3K4me2 or antagonistic H3K27me3 (Polycomb-mediated) methylation marks, on expression of a family LTR retroelements, either as an artificial COPIA93 LTR-GUS reporter transgene or as an endogenous element (COPIA93 LTR) in Arabidopsis. They further test how the methylation status relates to the gene-regulatory function of another endogenous LTR copy (solo LTR) which lies in the promoter of a pathogen recognition gene (RPP4). Key results which are well supported by the data are: i) a low-methylated LTR-GUS transgene behaves like an immune-responsive gene after a bacterial PAMP activation, using known immune-responsive WRKY TF-binding promoter elements for PAMP-induced expression, ii) DNA-methylation and associated H3K4me2 marks on an endogenous retroelement (AtCOPIA93) coexist and operate in concert with Polycomb-directed alternative H3K27me3 to restrict TN expression in response to a PAMP. This suggests that expression of certain TNs in the genome uses back-up methylation systems. iii)

endogenous solo-LTR in the RPP4 promoter, through its epigenetic (low-methylated) character and W-boxes, allows higher PAMP responsiveness of the RPP4 gene, and thus potentially an advantageous mode of immunity gene regulation in pathogen resistance. Overall, the ms is well written and the genetic, ChIP and methylation profiling experiments carefully and creatively done. The data here will be of general interest to immunity/stress and evolutionary biologists. I have a few comments:
1. Authors argue that that the TN chromatin status confers on nearby genes a particular mode of expression. Am wondering how general this is for immunity or stress-responsive genes. We would expect to see a statistical correlation between TN-methylation status and stress responsiveness/modes of expression of genes in the genome if this is a general trend. Or is this too simplistic?
2. Is the solo-LTR in RPP4 promoter an unusual case or do authors think it is generally applicable to immunity gene control? Am trying to get at whether the mode of RPP4 regulation is distinct from the majority of PAMP-responsive genes which don't have TN-LTR insertions near or in the promoter. Does methylation of these gene promoters limit their induced expression in PTI?
3. Is TN-insertion into protein-encoding gene regions associated with removal of methylation marks on those genes, such as seen in Solo-LTR, or is this an unusual case? The endogenous COPIA93 LTR only responds to PAMP when methylation is depleted. The methylation status is likely important to limit expression of dangerous genes, so the solo-LTR/RPP4 combination is potentially dangerous. It's not clear what determines these different epigenetic states of the LTR TNs - worth some discussion.
4. In Fig. 5, authors provide compelling evidence that Solo-LTR in the RPP4 promoter increases RPP4 basal and PAMP-triggered transcript accumulation. Do the RPP4 wt and delta LTR/W-box lines show different levels of bacterial or Hpa growth after PAMP trigger? ie - does LTR-control of RPP4 expression have a biological consequence for host and pathogen? If it is one of several of many similarly-controlled NLR genes, their up-regulation in PTI is likely to prime the immune system.
5. In Fig. 5 authors argue that induced response of LTR-GUS plants to Hpa-derived PAMP NLP20 indicates relevance of TN-mediated RPP4 regulation to Hpa. Why didn't they test the RPP4 transgenic lines to make this point?

1st Revision - authors' response 16th Febuary 2018

We would like to thank the reviewers for their constructive comments and for giving us the opportunity to further reflect on certain aspects of our study. We have addressed each suggestion to the best of our ability. Reviewer comments are below in **bold type** and our response is in regular type.

Referee #1:

The authors of the manuscript "Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons" describe the involvement of the retrotransposon AtCOPIA93 (EVD) in the response to biotic stress. By using an LTR::GUS reporter line, the authors show that this element is transcriptionally regulated upon bacterial and elicitor treatment. They furthermore show that the regulation of the endogenous element is controlled by DNA and PRC2-mediated H3K27me3 histone methylation. Interestingly, the LTR of these elements harbors cis regulatory motives that regulate the transcriptional activity of neighboring genes, including the resistance gene RPP4, which the authors investigate in more detail. This manuscript is a nice example of the effects of retrotransposon insertions on neighboring gene expression upon biotic stress and provides evidence that LTRs harbor cis-regulatory motifs important for regulation. I have some critical comments aiming to improve this manuscript:

1. It would be interesting to know whether the LTR::GUS reporter is also targeted by H3K27me3. Since PRC2 targeting is largely determined by cis elements, one would expect the transgene to become targeted by H3K27me3. If not, this would argue against a cis-element-

mediated PRC2 targeting of this LTR, which would be of interest in respect to currently unknown targeting mechanisms of TEs by PRC2.

This is an interesting question and we agree that this experiment brings some information on the mechanism at play for H3K27m3 deposition at *EVD*. We have now included a figure in the manuscript (Figure EV3E), and amended the text accordingly with the description of this result and its implications in the paragraph before last, page 7 and in the discussion page 14. As now stated and discussed in the MS, we could show that H3K27m3 was absent at the transgenic LTR sequence when performing H3K27m3-ChIP analysis on LTR::*GUS* transgenic plants, thus the transgenic LTR is not subjected to H3K27m3 methylation upon transformation. Therefore, the LTR sequence is unlikely to contain Polycomb response elements that recruit PRC2. One possible mechanism for H3K27m3 deposition at *ATCOPIA93-EVD* could be that a sequence other than the LTR (e.g. in the CDS) is responsible for recruiting PRC2 and for the nucleation of a H3K27m3 domain over the TE; in this scenario, one would expect more H3K27m3 on ATCOPIA93-*ATR* (which shares the same sequence as *EVD*)–unless it is strongly antagonized by higher levels of H3K9m2 than observed at *EVD*. However, we currently do not favor this hypothesis. Instead, we propose that the specific genomic context of *EVD* (embedded in a H3K27m3-rich region) and the almost complete loss of H3K27m3 at *EVD-LTR* in plants mutated for CLF –which was recently shown to be involved of the spreading phase of H3K27m3 at the *Flowering Locus C* (*FLC*) (Yang et al., 2017) – rather support the hypothesis wherein H3K27 trimethylation at *EVD* is favored by spreading from the neighboring genes marked by H3K27m3. As mentioned in the discussion, future studies should allow us to identify and delineate TEs that can potentially recruit PRC2 in *cis* (and through which sequences) and the ones that become H3K27m3 methylated because of spreading of this mark beyond a nearby genic nucleation site.

2. The relevance of the double layer of regulation of EVD by DNA methylation and H3K27me3 has not been addressed. Investigating expression of EVD in clf met1 or clf ddm1 double mutants would be helpful to address this question.

We agree that this is an important point and in order to address this question, we have generated a double *ddm1 clf* mutant. Expression analyses of *ATCOPIA93* in this background are now presented in Fig 4C, 4D (and Fig EV4E) showing the relevance of the double-layer of *EVD* regulation by DNA methylation and Polycomb (representative pictures of the newly generated mutant are also presented in Fig EV4D). These results are now described in the text page 9 as follows:
"To address the relevance of the double layer of regulation of EVD by DNA methylation and H3K27m3, we generated a double ddm1 clf double mutant (Fig EV4D). In ddm1 clf plants, even in the absence of elicitation, ATCOPIA93 mRNA levels were higher, in average, than in any other genetic background analyzed; in addition, we could observe a significant and quantitatively important increase of ATCOPIA93 expression upon PtoΔ28E challenge (Fig 4C) compared to the wild type and the single mutants. Although it is possible that this striking synergistic effect of ddm1 and clf mutations is contributed by indirect effects, these results show that ATCOPIA93 expression is dually controlled by both DNA methylation and Polycomb-mediated silencing. Pyrosequencing of the cDNA further showed that it was mostly EVD that was reactivated in clf- and in ddm1 clf-elicited mutant backgrounds (Fig 4D, Fig EV4E). This indicates that in the absence of silencing marks EVD tends to be more transcribed or expressed than ATR, possibly because of position effects, and this balance is shifted towards one TE or the other depending on the present marks in each genetic background."

We thus thank the reviewer for giving us the opportunity to reveal these striking molecular phenotypes in the double-mutant.

3. It would strongly increase the value of the manuscript if the authors could test whether PAMP elicitation does induce new insertions of EVD. Since the authors have the material, doing a transposon display should not be that difficult to address this question.

We agree and have tried to answer this question to the best of our abilities. As now stated in page 9, somatic transpositions in adult leaves upon bacterial elicitation, even if they occurred, would unlikely be detected by this technique. In fact, new insertions would be unique as there is no cell

division at this stage, thus they would not be clonally inherited nor could not be detected at later stages of development, since adult leaves -contrary to the shoot apical meristem- do no contribute to the germ line. Cloning and detecting one single insertion in one cell by PCR would be extremely technically challenging. We have thought of repeating this experiment in young seedlings, however at this stage, the LTR::*GUS* is not well induced during PTI (Fig EV4A) thus it is not a relevant stage to perform this experiment. In the future, we would like to perform elicitation of the tissues (flowers) that give rise to the germline in order to identify in the progeny individuals that display new insertions in their genome. However, these experiments require significant set-up (elicitation methods in flowers, prediction of the number of transposition events to anticipate the number of seedlings to be tested in the progeny to detect one insertion, etc.) and should be the focus of another study to address specifically this exciting question of enhanced transposition upon bacterial elicitation.
Nevertheless, to try to address the referee's request, we attempted to detect intermediates of transposition in the form of extrachromosomal, linear *ATCOPIA* DNA (ecDNA) in *ddm1 clf*, and we could indeed visualize them in this double mutant background. These explanations and results are now presented in Fig4E and are described in the text page 9 as follows:
"We next wanted to test whether the high levels of ATCOPIA93 expression in ddm1 clf mutants upon bacterial elicitation could lead to ATCOPIA93 transposition. Detection of transposition by southern blot and transposon display requires the new insertions to have been clonally inherited during cell division; therefore, these two techniques are not sensitive enough to detect transposition events upon bacterial elicitation of adult leaves, a developmental stage in which cell division has mostly ceased. Alternatively, we thought of detecting intermediates of transposition in the form of retrotranscribed, linear, extra-chromosomal (ec)DNA. After ligating the linear extra-chromosomal DNA to adaptors (Takeda et al, 2001; Mirouze et al, 2009b), we could clearly detect, in the ddm1 clf mutant, ATCOPIA93 intermediates of transposition (Fig 4E, top panel)."

To further test for an increase of these forms upon bacterial stress, we attempted to quantify the cloned ecDNA forms by qPCR 24 hours post-infiltration. We saw for two biological replicates out of three an increase of *AtCOPIA93* ecDNA upon elicitation. Predictably, this increase is lower than the one observed for *AtCOPIA93* mRNA in *ddm1 clf* upon bacterial elicitation (Fig 4C), presumably due to the reported excess of the alternative GAG-coding RNA isoform over the full-length transcript *(Oberlin et al., 2017)*. In sum, while transposition detection remains a goal, we can at least propose that *ATCOPIA93* potential for transposition increases during innate immune response in the *ddm1 clf* background. These results are presented in the bottom panel of Fig 4E, and commented in the same paragraph page 9.

4. Data shown in Fig. 1B, it is unclear how many lines were analyzed.

In Fig 1B, only one homozygous transgenic line is analyzed and presented: we have now specified this in the legend. However, a total of three independent lines were selected on the basis of 3:1 segregation ratios of the transgene (single insertion) and brought to T3 generation where the transgene was in a homozygous state. All three lines behave similarly as for GUS expression 24hpi after bacterial stress and lack of DNA methylation: the Figure EV1 has now been amended to include these data (Fig EV1B and D). Fig EV1 now also includes additional bisulfite analyses of the LTR transgene (Fig EV1B) in one additional line.

5. In figures 2B, 4A and 4B the variation between the experiments is quite high. A statistical treatment should be performed to test whether the observed differences are significant.

We have indeed been aware of this variability between experiments and acknowledge it. First, we would like to comment on it. Although we have attempted to reduce it as much as possible by eliciting the plants at the same age and time of the day and trying to reproduce the same environmental conditions such as temperature and hygrometry, we could not avoid variations which are inherent to elicitation experiments in adult leaves. We found that these variations in the homogeneity, extent and mostly timing of the PTI response were unavoidable in adult leaves. We have now amended the legends of Figure 2 and 4 with a concise version of this explanation for the reader:

"The variability observed between biological replicates is inherent to the developmental stage analyzed (adult leaves) which often shows differences from one experiment to the other, in the extent and timing of PTI (see Fig 1C/EV1E and Fig2A/EV2)."

Given this inter-experiment variability and to address both the referee and editor's request, we performed paired T-tests to assess the significance of our results. In addition, we have added additional biological replicates (one additional independent experiment for each panel in Fig 2B and two additional experiments for Fig 4A) to the analyses. We could see reproducible and significant differences in *AtCOPIA93* expression between wild-type and DNA methylation mutants upon bacterial stress (Fig2B and 4A). In the PRC2-mutant *clf*, we initially saw a stronger induction of *ATCOPIA93* compared to wild-type two times out of three. However, with the additional replicates produced during the revisions, this observation was made two times out of six and the increased activation compared to the wild type was no longer significant, despite a trend towards enhanced EVD mRNA levels in the *clf*-elicited background. This is possibly due to variations in temperature and floral transition as we noticed that, despite our controlled conditions, the *clf* mutant flowered a bit later in these new experiments; in addition, the concomitant activation in *ddm1* mutant, stronger than previously observed, points to a shift in the kinetics of the response in these newly generated datasets. This result can nonetheless be understood now in the light of the strong reactivation of *ATCOPIA93* in *ddm1 clf* (now Fig 4C) supporting functional redundancy between Polycomb and DNA methylation.
We now amended the text page 8 as follows:
"Results from these analyses revealed a modest increase of ATCOPIA93 expression in bacteria-elicited clf plants compared to the wild type plants, though less reproducible and weaker than in ddm1-elicited plants (Fig 4A, EV4A)."

6. In figure 3B, the genes analyzed are not mentioned in the main text.

This has now been amended in the manuscript in the corresponding section page 6.

7. In Fig. 3A the authors show data of ChIP-Chip experiments. In the Supplement, they should include data of published ChIP seq experiments (for wild-type), to add additional support for the presence of H3K27me3 on the LTR of EVD.

To reply to the request of the reviewer, we recovered the raw H3K27m3 ChIP-seq data from *Wang et al, 2016* (wild-type seedlings). We could confirm what we saw in ChIP-chip data: H3K27m3 is present at the *EVD* LTR sequences. In addition, we remapped their H3K27m3 ChIP-seq data from *clf* seedlings, and consistent with our ChIP-qPCR results, observed the loss of H3K27m3 marks at *EVD* in this background. We have now included these data in Fig EV3A (Fig EV3B) and amended the text accordingly.

8. Fig. 3D: It unclear what exactly is plotted and what the points refer to.

This has now been clarified in the corresponding legend with the following amendment: *"The values plotted correspond to the ratio between the amount of amplified DNA in the Sau96I digestion and the amount of amplified DNA in the undigested control, as calculated by the formula* $2^{-(Ct.digestedDNA-Ct.undigestedDNA)}$ *and using primers specific for a region of EVD LTR spanning this Sau96I restriction site. Dark and light symbols are used for the first and second experiments respectively."*

9. I find it confusing that the authors use the name EVD and ATCOPIA93 synonymously. It would improve clarity to only use the name EVD as it has been previously published. It is standard in the field to use the gene name that has been published first; the same rule should be applied here for a TE.

Our intention was: by "*ATCOPIA93",* we refer to the family of *ATCOPIA93* retrotransposons composed in particular of *EVD* and *ATR* (each corresponding to one single locus) and by "*ATCOPIA93-EVD*" we refer solely to *EVD*, but agree this can be confusing. For clarity, we have

now replaced "*ATCOPIA93-EVD*" by the published name of the locus "*EVD*" as suggested, throughout the text.

10. References are not correct:
-page 2: "The maintenance of CHH methylation..." Deleris et al., do not address the mechanism of CHH methylation; Zemach et al., 2013 should be correctly cited here.
We have now removed this citation and replaced it by "Zemach et al., 2013" as well as "Stroud et al, 2013". Thank you for having noticed that.

-page 6: Ordonnez et al., 2014 is not present in the reference list and also not the correct reference in this context.
We have now added in the text the appropriate specific references showing genome wide losses of CG methylation in *met1* (Cokus et al., 2008, Lister et al., 2008) and *ddm1* (Stroud et al., 2013) and chose to place them at a more appropriate location *i.e*. in the introduction Page 2.

11. Page 7, rephrase sentence :"...EVD almost exclusively is reactivated..."to "it was mostly EVD that became reactivated in the clf-elicited background."

We thank the reviewer for suggesting this clearer alternative; the text has now been modified accordingly.

12. Page 8, rephrase "---detection of transcription was detected..."

We thank the reviewer for spotting this mistake; this sentence has now been replaced by *"Interestingly, transcription was detected in response to various bacterial challenges downstream of soloLTR-1, soloLTR-2 and soloLTR-5 (Fig EV5A)."*

Referee #2:

The manuscript by Deleris et al., "Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons", describes the identification and functional characterization of a TE-derived pathogen-responsive promoter in Arabidopsis. Transgenic reporters reveal transcriptional activation of the LTR of a Copia element upon exposure to several pathogens, depending on lack of epigenetic silencing components and the presence of W box domains. The responsiveness seems to be determined by DNA methylation and Polycomb-based histone methylation. A soloLTR derived from the TE was found in the promoter of a pathogen-responsive gene.
These findings are interesting, and the work is largely well designed and performed. It provides another example for the potential role of TEs as regulatory elements, and it will be of interest to a broader readership. There are a few points that should be discussed.

1. Reproducibility between experiments is poor, as becomes evident from the lack of error bars in expression and ChIP experiments. However, the authors provide replicates as separate figures (e.g. 2A and EV2), but these often show substantial differences, in this case in kinetics as well as in the degree of response. Is this classical for pathogen infection experiments?

With regards to the ChIP experiments, we admit that the amount of starting material (described in the materials and methods) for chromatin extraction/IP was not always consistent for every independent experiment due to plant material availability, which may have impacted ChIP efficiency. In addition, we were forced to use multiple batches of antibody, which also could have contributed to variance. However, the results we describe can be observed in every experiment, thus are reproducible. We thank the referee for having noted that and his/her understanding in this respect.

As for the variability of expression data upon bacterial treatment, another referee also brought up this point. We now have added more replicates and performed statistical tests, and our results,

despite being variable, are significant and reproducible, with the exception of the single *clf* mutant (addressed in our comments to Referee #1). However, we agree that the variability observed deserves explanations that should also be provided to the readers.
As explained to Referee 1, we have attempted to reduce variability as much as possible by eliciting the plants in environmentally-controlled conditions; however, we could not avoid variations which are indeed inherent to pathogen infection/elicitation experiments in adult leaves. Indeed, these variations were unavoidable with respect to the homogeneity, extent and timing of the PTI response in adult leaves. These variations can certainly be reduced in seedlings as can be seen in the Figure EV4A, but the LTR::GUS construct is more lowly expressed in these tissues.
We have now amended the legends of Figure 2 and 4 with a concise version of this explanation:
"*The variability observed between biological replicates is inherent to the developmental stage (adult leaves) analyzed which often shows differences from one experiment to the other, in the extent and timing of PTI (see Fig 1C/ EV1E and Fig 2A/EV2)".*

2. To judge the role of DNA methylation at the EVD LTR, the number of Cs in the different contexts should be shown. The information can be extracted from Figure EV1A but a close look reveals that (1) the first 2 of 7 CG positions in Figure 1 A are not within the region marked as LTR (and the Discussion mentions 5 sites, there are also only 6 CHG sites), and (2) less than half of all CHH sites are methylated at all, and most in just 1 of the 17 clones analysed. Other Copia LTRs are also characterized by low CG/CHG frequencies, and this special feature should be better discussed.

We have now indicated the number of C in different contexts at the LTR sequence on the side. In addition, each C that is not methylated at all is now represented by a dot so that the reader can understand more easily that only Cs are depicted on the graphs, and as the reviewer notes, less than half of CHH are methylated at all at the endogene. Finally, for clarity, we only show the LTR sequence (and removed any sequence analyzed upstream of it).

We thank the referee for pointing out that other Copia LTRs are characterized by low CH/CHG frequencies. We have now ourselves analyzed the count of CG and CHG di- and trinucleotides/100 bp respectively at all Arabidopsis Copia LTRs and compared them with Gypsy LTRs as well as to other genome compartments (exons, introns, UTRs, TEs). Indeed, we found significantly less CG and CHG sites at Copia LTRs. These observations are now presented in Appendix Figure S1 and we have discussed them in page 13 within the paragraph below:
In spite of the reported antagonist effect of DNA methylation- in particular CG methylation- on H3K27m3 deposition, we found that the two marks could co-occur at the EVD 5'LTR in vegetative tissues. To explain this observation, we propose that, like in mammals, a low density of CG methylated sites is permissive to the deposition of H3K27m3, since EVD LTR sequence contains only five CGs (i.e. a frequency of 1,2/100bp) and 6 CHG (i.e. a frequency of 1,4/100bp). Interestingly, other Copia LTRs are also characterized by low CG and CHG frequencies compared to other genomes segments and we observed that, in average, the CG and CHG density at Copia LTRs is significantly lower than at Gypsy LTRs (Appendix Figure S1). The first implication of this low CG and CHG density for the LTRs of Copia elements could be that DNA methylation has a less repressive potential at Copias than at other TEs, supporting the idea that epigenetic control of TE silencing should be studied by taking into account TE class (Underwood et al, 2017). On the other hand, this low density of CG methylation could be mechanistically linked to an increased permissiveness to H3K27m3 deposition (Statham et al, 2012; Brinkman et al, 2012) and Copias LTRs could thus be more prone to be targeted by H3K27m3 which would compensate for less repression by DNA methylation.

We thank the reviewer for giving us the opportunity to reflect on this interesting aspect and to discuss it in our paper.

3. The analysis of just one transgenic line by bisulfite sequencing is unsatisfying. The added semi-quantitative analysis at one CG position by restriction digest analysis in Figure EV1B indicates no substantial methylation at the transgene also in other lines, but these data are weak.

We have now analyzed one more T3 line with bisulfite sequencing (LTR::GUS#6) and these analyses are now presented in Figure EV1B. This panel now also includes additional bisulfite analyses of the LTR transgene in the next (T4) generation of the LTR::GUS line#12.

4. H3K27me3 analysis: why was the transgenic line not included? As it was not DNA methylated (at least in the line analysed), it would have been informative to ask whether the LTR attracts the histone methylation independent of DNA methylation. If so, it could have revealed whether and how this modification eventually changes upon infiltration of flg22 or Ptodelta28E. This could have complemented the comparison between wt and clf, especially as clf leaves are morphologically quite different, to begin with, whereas the transgenic line has a normal leaf phenotype.

This is an interesting experiment (also suggested by Referee 1) and we now have included the corresponding figure in the manuscript (Figure EV3E) and amended the text accordingly.
As now stated and discussed Page 7 and 14, we could show that H3K27m3 was absent at the transgenic LTR sequence when performing ChIP analysis on LTR::*GUS* transgenic plants. Nevertheless, this negative result brings some information on the mechanism at play for deposition of H3K27m3 at *EVD*, which we have now discussed further in the corresponding section page 7 and discussion page 14.

5. Role of the soloLTR5: it would have been much more elegant to eliminate the LTR from the RPP4 gene with gene editing, to avoid position effects in different transgenic lines. However, I accept the data from pools of primary transformants with ectopic insertions of the two different constructs in the rpp4 mutant background and their significant difference (although the different pool sizes in Fig. 5C, 33 versus 18, weaken the statistics).

We agree with the reviewer that genome editing would have been a superior approach, but this is not for lack of trying: we attempted twice to eliminate the soloLTR5 by CRISPR; we found out that our second attempt was unsuccessful during the revision process and did not have time to give it another try.
We thank the reviewer for acknowledging the validity of our alternative method. As for the difference in pool sizes, it is due to the number of primary transformants available at the time. We did not want to add more "pRPP4ΔLTR:RPP4" individuals in an independent experiment during the revisions to avoid creating a bias in the results. Instead, we performed the statistical test by sampling randomly 18 "pRPP4:RPP4" transformants out of the 33, four times. As for the Mock *versus* PtoΔ28E treatments, as predicted by the reviewer, the p-value increased but, importantly, we still obtained significant induction in the pWT:RPP4 plants for all samplings (p value < 0.02 in all four cases, see below) while no significant induction was observed in the pΔLTR::RPP4 plants. We thus decided to present the results with one representative sampling in Fig 5C.

Table Analyzed	20180201RPP43sample18-A
Column C	pWT::RPP4 + PstΔ28
vs.	vs.
Column A	pWT::RPP4 + mock
Unpaired t test with Welch's correction	
P value	0.0201
P value summary	*
Significantly different? (P < 0.05)	Yes
One- or two-tailed P value?	Two-tailed
Welch-corrected t, df	t=2.513 df=21.34

Table Analyzed	20180201RPP43sample18-B
Column C	pWT::RPP4 + PstΔ28
vs.	vs.
Column A	pWT::RPP4 + mock
Unpaired t test with Welch's correction	
P value	0.0165
P value summary	*
Significantly different? (P < 0.05)	Yes
One- or two-tailed P value?	Two-tailed
Welch-corrected t, df	t=2.591 df=22.43

Table Analyzed	2080201RPP43sample18-C
Column C	pWT::RPP4 + PstΔ28
vs.	vs.
Column A	pWT::RPP4 + mock
Unpaired t test with Welch's correction	
P value	0.0193
P value summary	*
Significantly different? (P < 0.05)	Yes
One- or two-tailed P value?	Two-tailed
Welch-corrected t, df	t=2.495 df=25.90

Table Analyzed	20180201RPP43sample18-D
Column C	pWT::RPP4 + PstΔ28
vs.	vs.
Column A	pWT::RPP4 + mock
Unpaired t test with Welch's correction	
P value	0.0148
P value summary	*
Significantly different? (P < 0.05)	Yes
One- or two-tailed P value?	Two-tailed
Welch-corrected t, df	t=2.638 df=22.54

6. The analysis of the role of the W box motifs in the LTR is inconsistent: while deletion of W2 in the GUS reporter seems to reduce the responsiveness more than that of W1 (Fig. 1E), only W1 was deleted in the soloLTR of the RPP4 construct (Fig. 5 B and D). Why?

We should have stated more clearly that the canonical W2 box is not present in the soloLTR of *RPP4*. This sequence difference can be extracted from figure EV5 but we did not emphasize it. We agree that this difference with the LTR of the full-length *EVD/ATR* should be stated in the text (and emphasized on the Fig EV5C) and this evolution of the soloLTR commented. We have now done so in the appropriate result section and discussion (page 10 and 15). The text has been amended with the two following parts:

"In addition, we wanted to assess whether soloLTR-5 contributes to RPP4 induction during PtoΔ28E elicitation through the W-box motif. We noticed that the W-box 2 present in EVD/ATR-LTR is absent in the soloLTR-5 as there is a single nucleotide polymorphism (SNP) in the core motif (GTCA>GTCG) (Fig EV5C). We thus tested the importance of the conserved W-box 1, which we had found earlier to have a partial effect on the induction of the LTR::GUS fusion (Fig 1E)"….

"This possible selection of an unmethylated soloLTR –combined with the loss of one of the two W-box-, could thus be seen as part of a RPP4 promoter maturation process towards well-balanced cis-regulation since sufficient expression of disease resistance genes is required to mediate immunity but their overexpression results in significant fitness costs (Lai and Eulgem 2017)".
This is actually in line with the comment of referee 3 "The methylation status is likely important to limit expression of dangerous genes, so the solo-LTR/RPP4 combination is potentially dangerous."

We thank both the reviewers for giving us the opportunity to reflect on this aspect.

7. P. 4 line 8 from bottom: "instead, we generated.." is misleading, make it clear that you just re-defined the purpose of the reporter line described in the first sentence of the paragraph.

We thank the reviewer for pointing out this confusing messaging. We have now amended the text as below:

"…initially to serve as a reporter of DNA methylation levels as previously reported for a Gypsy-type LTR-retrotransposon (Yu et al., 2013). Unexpectedly, the LTR::GUS transgenes were not methylated in any of the transgenic lines obtained (Fig 1A, Fig EV1B). Thus, instead, we used the LTR::GUS as a reporter of ATCOPIA93 promoter activity that we could exploit to assess ATCOPIA93 responsiveness during PTI in the absence of DNA methylation-mediated control."

8. Coexistence and cooperation of H3K27 methylation and DNA methylation in repression is not as novel as claimed here. A lot of literature is available in cancer epigenetics and should at least be mentioned.

It is true that coexistence has been well reported in mammals, and to underscore this, we moved the corresponding references from the discussion to the result section page.
"This co-occurrence of the two marks, while unexpected and never reported in vegetative tissues, was recently observed in the endosperm at pericentromeric transposable elements (Moreno-Romero et al, 2016) as well as in mammals where lower densities of CG methylation were found to allow H3K27m3 deposition (Statham et al, 2012; Brinkman et al, 2012)."

As for cooperation, we have now added in the discussion (page 15) what seemed to us the most relevant references for evidences of cooperation in the form of functional redundancy : in naturally hypomethylated plant cell types such as the plant endosperm (Weinhofer *et al*, 2010) or chemically-hypomethylated mammals ESCs (Walter *et al*, 2016).

9. More care could be taken to cite the correct primary references. Just one example: impaired CG methylation in met1 mutants was described prior to Ordonez et al. 2013.

We thank the referee for pointing out this confusing reference (also pointed out by Referee 1) that now has been removed. We have now added in the text the appropriate specific references showing genome-wide losses of CG methylation in *met1* (Cokus et al., 2008, Lister et al., 2008) and *ddm1* (Stroud et al., 2013) and chose to place them in the introduction Page 2.

10. Provide a reference for the Tanaka method in Material and Methods.

This reference (Hurkman and Tanaka, 1986) has now been added to the Material and Methods.

11. Editing of the manuscript by a native English speaker could improve grammar and punctuation.

The manuscript has now been read by a native English speaker.

Referee #3:

This analysis provides evidence of a causal role of transposons in regulation of plant immunity gene expression. Previous studies have shown that TN mobility is a driver of genome evolution and that TNs can be mobilized in response to various stresses. There is also evidence that TN insertions in the genome attract specific epigenetic marks to limit their activity and these marks can spread to cis-regulation of other genes. Here, the authors examine the effect of the epigenetic environment - DNA methylation and associated H3K9me2 or antagonistic H3K27me3 (Polycomb-mediated) methylation marks, on expression of a family LTR retroelements, either as an artificial COPIA93 LTR-GUS reporter transgene or as an endogenous element (COPIA93 LTR) in Arabidopsis. They further test how the methylation status relates to the gene-regulatory function of another endogenous LTR copy (solo LTR) which lies in the promoter of a pathogen recognition gene (RPP4). Key results which are well supported by the data are: i) a low-methylated LTR-GUS transgene behaves like an immune-responsive gene after a bacterial PAMP activation, using known immune-responsive WRKY TF-binding promoter elements for PAMP-induced expression, ii) DNA-methylation and associated H3K9me2 marks on an endogenous retroelement (AtCOPIA93) coexist and operate in concert with Polycomb-directed alternative H3K27me3 to restrict TN expression in response to a PAMP. This suggests that expression of certain TNs in the genome uses back-up methylation systems. iii) endogenous solo-LTR in the RPP4 promoter, through its epigenetic (low-methylated) character and W-boxes, allows higher PAMP responsiveness of the RPP4 gene, and thus potentially an advantageous mode of immunity gene regulation in pathogen resistance. Overall, the ms is well written and the genetic, ChIP and methylation profiling experiments carefully and creatively done. The data here will be of general interest to immunity/stress and evolutionary biologists. I have a few comments:

1. Authors argue that that the TN chromatin status confers on nearby genes a particular mode of expression. Am wondering how general this is for immunity or stress-responsive genes. We would expect to see a statistical correlation between TN-methylation status and stress responsiveness/modes of expression of genes in the genome if this is a general trend. Or is this too simplistic?

We thank the reviewer for giving us the opportunity to assess how general could be our findings with the comments 1, 2 and 3. To address these comments, we performed genome-wide analyzes now presented in Figure S2 (Appendix) and amended the discussion with the following (Page 15).

"Previous genome-wide studies have suggested that methylated TEs have a genome-wide repressive effect on nearby gene expression -correlated with their proximity to the gene- (Hollister et al, 2011; Wang et al, 2013) and also that, upon pathogen stress, transcription of immune-responsive genes could be coupled to the dynamic methylation state of the proximal TEs (Dowen et al, 2012). Whether the selection/presence of a constitutively unmethylated TE sequence, a fortiori a soloLTR which contains cis-regulatory motives, in the promoter of PAMP-responsive genes is a general mechanism that contributes to innate immunity is an exciting question which deserves further exploration. In a first attempt to generalize our findings, we searched for and found dozens of immune-responsive genes whose promoters overlap with a TE fragment containing one LTR or an LTR-derived sequence (Appendix Figure S2A and Table S2). Interestingly, most of them (67%) were unmethylated (Appendix Figure S2B). Furthermore, it is that last group that comprised the genes that were the most induced by PAMPs and they were also more induced in average, compared to the group with a methylated LTR in their promoter (Appendix Figure S2C). It would certainly be too simplistic to draw correlations between TE methylation status and downstream gene induction during immune response as many other parameters have to be taken into account: presence/absence of cis-regulatory elements in the TE, strength of the promoter, possible differential impact of different methylation contexts on transcription, localization of DNA methylation within the promoter) and the number of genes too low; nonetheless, the genes we have identified are potential candidates for further studies to generalize the phenomenon we have uncovered at RPP4. These studies will certainly benefit from currently arising techniques such as epigenome editing which

should allow in the future to methylate or demethylate specifically a particular/discrete TE to test for its impact on nearby gene transcription."

2. Is the solo-LTR in RPP4 promoter an unusual case or do authors think it is generally applicable to immunity gene control?
Am trying to get at whether the mode of RPP4 regulation is distinct from the majority of PAMP-responsive genes which don't have TN-LTR insertions near or in the promoter.

Our response can be found in the previous point (1.), discussion and Figure S1. Briefly, we found that this is not an unusual case and that about 1% of PAMP-induced genes contain a LTR or LTR-derived sequence overlapping with their promoter (defined as 1KB upstream the TSS, see description of the methods in the Appendix).
With regards to the generality of this type of control, we would also like to share with the reviewer this interesting parallel in humans: an endogenous retrovirus (ERV)-derived sequence was recently found to be required for the expression of a couple of Interferon (IFN)-induced genes and proper immune response against viral infection (Chuong et al., 2016). Thus, in both human and plant lineages, retroelements/retroviruses sequences have been coopted for proper expression of genes involved in immune response.

3. Is TN-insertion into protein-encoding gene regions associated with removal of methylation marks on those genes, such as seen in Solo-LTR, or is this an unusual case? The endogenous COPIA93 LTR only responds to PAMP when methylation is depleted. The methylation status is likely important to limit expression of dangerous genes, so the solo-LTR/RPP4 combination is potentially dangerous. It's not clear what determines these different epigenetic states of the LTR TNs - worth some discussion.

We have answered these questions one by one below:

Is TN-insertion into protein-encoding gene regions associated with removal of methylation marks on those genes, such as seen in Solo-LTR, or is this an unusual case?

Our response can be found in the previous point (1.), discussion and Figure S1. Briefly, among the 1% of PAMP-induced genes that contain an LTR or LTR-derived sequence in their promoter that we have identified (see above), we found that for 67% the LTR was unmethylated (with 30% methylated, and 3% with no information). In addition, among the PTI-induced genes containing a TE<1kb in their promoter (control), 67% had an unmethylated TE in their promoter (16% had both a methylated and unmethylated TE). These seem high percentages given that TEs are usually thought to be DNA methylated. Also, this high proportion of unmethylated sequences does not seem specific to soloLTRs and LTR-containing sequences: instead, this might have to do with a "relic/size" effect of these TE<1kb that we selected to address the question, and possibly with position effects. We thank the reviewer for giving us the opportunity to make these interesting observations that deserve further exploration.

The endogenous COPIA93 LTR only responds to PAMP when methylation is depleted. The methylation status is likely important to limit expression of dangerous genes, so the solo-LTR/RPP4 combination is potentially dangerous.
It's not clear what determines these different epigenetic states of the LTR TNs - worth some discussion.

We thank the reviewer for giving us the opportunity to discuss this aspect. Mechanisms for negative control (notably, *RPP4* is subjected to multiple levels of negative control including PTGS (Yi et al., 2007)) are thought to have evolved to reduce fitness costs associated with resistance gene expression; on the other hand, transient expression of disease resistance genes is required to mediate immunity. We thus propose that the positive selection of an unmethylated soloLTR (and later on the loss of the W-box2) could be part of a *RPP4* promoter maturation process towards well-balanced *cis*-regulation, as discussed with Referee 2 and now stated in the discussion.

As for the basis for the unmethylated status of the soloLTR, we had started to discuss this page 15 and have now further expanded this paragraph:

"This lack of de novo methylation upon transformation may be explained by weak LTR transcriptional activity in untreated plants thus preventing expression-dependent RNA-directed DNA Methylation (Fultz & Slotkin, 2017) and/or by low levels of CHH methylation/siRNAs at ATCOPIA93 (Fig 1A) (Mirouze et al, 2009; Marí-Ordóñez et al, 2013)), thus preventing identity-based silencing in trans (Fultz & Slotkin, 2017). Similarly, in wild type plants, the ATCOPIA93-soloLTRs appear usually to be constitutively unmethylated (Fig EV5A), in particular the soloLTR-5 described in detail here…
…We propose that the low levels of EVD LTR siRNAs, which could methylate the soloLTR-5 in trans if present in larger quantities, could be the result of evolutionarily selection, so that soloLTR-5 remains unmethylated and proper immune response can be properly activated…
… This possible selection of an unmethylated soloLTR –combined with the loss of one of the two W-box-, could thus be seen as part of a RPP4 promoter maturation process towards well-balanced cis-regulation since sufficient expression of disease resistance genes is required to mediate immunity but their overexpression results in significant fitness costs (Lai and Eulgem 2017)."

4. In Fig. 5, authors provide compelling evidence that Solo-LTR in the RPP4 promoter increases RPP4 basal and PAMP-triggered transcript accumulation. Do the RPP4 wt and delta LTR/W-box lines show different levels of bacterial or Hpa growth after PAMP trigger? ie - does LTR-control of RPP4 expression have a biological consequence for host and pathogen? If it is one of several of many similarly-controlled NLR genes, their up-regulation in PTI is likely to prime the immune system.

First, we would like to clarify that, even if it is probable that other NLRs are similarly regulated through the presence of a TE-derived *cis*-regulatory motives in their promoter (based on the candidates now established in Fig S1), *RPP4* is the sole NLR gene regulated by a *COPIA93* soloLTR. Since we only deleted/modified the soloLTR5 element in the ΔLTR/w1-box lines, only the expression of RPP4 is impacted.
There is a growing body of evidence (Boccara et al., 2014, Bonardi et al., 2011) that genes involved in race-specific resistance are also involved in PTI. However, the impact on PTI that has been reported is through mis-regulation of a subset of Coiled-coil NBS-LRR (CNL) genes coordinately targeted by a microRNA. Thus, priming is likely to be mediated by the coordinated overexpression of multiple CNLs. Consequently, it seems unlikely that the mis-regulation of only one NBS-LRR gene, *RPP4*, would cause priming. Nevertheless, we tested whether a mutant for *RPP4* could impact the primed state induced by PTI by looking at one outcome of early immune responses, namely the cell wall reinforcement by callose deposition (Boller and Felix, 2009). We observed no significant decrease in *rpp4* mutant *versus* WT (Col-0) in response to either flg22 or PtoΔ28E, showing that the absence of RPP4 transcripts–let alone their downregulation by deletion of the soloLTR5–is not sufficient to impair the outcome of PTI.
Consequently, we did not do this particular experiment of testing whether there is less PAMP-induced priming in the ΔLTR /w-box RPP4 line.

However, we agree that it could have been very interesting to answer the broad question of the referee "**ie - does LTR-control of *RPP4* expression have a biological consequence for host and pathogen?**," through another experiment.
In fact, given that *RPP4* mediates Race specific resistance to the oomycete *Hpa* race Emwa, we thought to test the functional relevance of the soloLTR-mediated upregulation of RPP4 transcripts levels by PAMPs for the compatible *RPP4*-Hpa Emwa interaction, which leads to full resistance to this oomycete strain. However, the feasibility of the experiment has been compromised by our experimental system. Despite the fact that we transformed the *rpp4* mutant and that the transgenic *RPP4* was under the control of its own promoter, we found that the RPP4 transcripts levels in the primary transformants were usually higher than endogenous RPP4 transcript levels in Col-0 wild type plants. In ΔLTR/w1-box mutant lines, even if the RPP4 transcripts levels were consistently lower compared to the pRPP4:*RPP4* lines, they were still in the same range as in Col-0 (see Fig 5C compared to Fig EV5B). Given that the endogenous levels of *R* genes involved in race-specific resistance already confer full-resistance to the pathogen, we thus could not test the hypothesis by which a decrease in *RPP4* expression causes less resistance in our system.

Alternatively, to address the relevance for this LTR-mediated regulation of *RPP4*, we performed the experiment suggested by the reviewer below.

5. In Fig. 5 authors argue that induced response of LTR-GUS plants to Hpa-derived PAMP NLP20 indicates relevance of TN-mediated RPP4 regulation to Hpa. Why didn't they test the RPP4 transgenic lines to make this point?

We now have included this additional experiment in the manuscript (Fig 5F) with the following description:
"Importantly, the induction of RPP4 in response to NLP20 was compromised when the soloLTR was absent or when the Wbox-1 was mutated (Fig 5F), showing that this regulation is relevant during immune response against oomycetes."

2nd Editorial Decision 20th March 2018

Thank you for submitting a revised version of your manuscript. It has now been seen by two of the original referees whose comments are shown below. As you will see, they both find that all criticisms have been sufficiently addressed and recommend the manuscript for publication, pending a few remaining text edits. While your manuscript is thus in principle accepted for publication in The EMBO Journal, I would ask you to submit a final revised version that includes the requested text changes as well as the following editorial points:

-> I noticed that Appendix fig S1 refers to an unpublished RNAseq dataset that is not included as part of the present study. We generally do not allow references to 'data not shown' in papers published in The EMBO Journal but since this analysis is used as a minor point in the discussion - not as a main argument in the results section - we can make an exception in this case. However, you should acknowledge this in the discussion when referring to Appendix fig S1 (specify that the RNAseq data derives from your own unpublished data).

-> The two Appendix tables need to be updated to fit the journal nomenclature. If you want to keep them as part of the Appendix then please insert them in the Appendix pdf (and include them in the table of contents). Alternatively, you can keep them in the excel format by making them Expanded View tables (please also update the callouts in the main manuscript file if you do so). See our online author guidelines for more detail http://emboj.embopress.org/authorguide.

-> I am not sure if the browser screenshot in fig 3A is of sufficiently high resolution to display well in the final version. Is it possible for you to include a version with either higher resolution or a clearer colour scheme?

-> The GFP-WB in figure 1D is extremely pixelated, I would suggest that you make the panel slightly smaller or use a higher resolution image for this

Thank you again for giving us the chance to consider your manuscript for The EMBO Journal, I look forward to receiving your final revision.

REFEREE REPORTS

Referee #2:

The manuscript by Deleris et al. has been substantially improved considering most of my previously raised points, and also much of the other reviewers' input. I admit that I am still a bit concerned about the high variability between experiments, but the authors have tried hard to reduce it and provide arguments why this is a difficult issue. I do not question the conclusions drawn on the data. I appreciate the much improved documentation of the methylation sites. I hope that the planned elimination of the solo-LTR will work in future and will strengthen the evidence for its role.

Referee #3:

This revised ms is an interesting and substantial piece of work. Authors have improved the data analysis and writing, and now place their findings more clearly in the context of other published work. Authors have responded satisfactorily to my specific comments on the importance and generality of LTR locations to immunity gene expression control, and Hpa infection experiments. Generally, the data are carefully reflected on, and I like the notion that LTRs might contribute to promoter maturation for balanced cis-control of immunity/stress-related genes. Inclusion of Fig. 5F showing responsiveness of RPP4 gene in the various rpp4 LTR/W1-box transgenics helps reinforce the PAMP effects on expression of this gene.
A few minor points:
1. I would remove words like 'interestingly' from text.
2. I don't see Waese et al (2017) and Ordonnez et al (2013) in refs - end 1st paragr. Pg 10. Can authors carefully check through references.
3. Figure S2 in Suppl. Materials is marked as Fig. S1. Can authors provide a short explanation to data in Appendix Tables S1 and S2.

2nd Revision - authors' response 13th April 2018

Editor/reviewer comments are below in **bold type** and our response is in regular type.

-> I noticed that Appendix fig S1 refers to an unpublished RNAseq dataset that is not included as part of the present study. We generally do not allow references to 'data not shown' in papers published in The EMBO Journal but since this analysis is used as a minor point in the discussion - not as a main argument in the results section - we can make an exception in this case. However, you should acknowledge this in the discussion when referring to Appendix fig S1 (specify that the RNAseq data derives from your own unpublished data).

This has now been acknowledged in the text p16
(*Appendix Figure S2C and Appendix Figure S3, where the data derive from our unpublished RNA seq data)*

We have also deposited the data on NCBI on the sequence read archive (SRA) and the accessions for the data submitted are described in the Appendix material and methods.
"The data are accessible on the NCBI sequence read archive (SRA) under accession (SRP133028) and a complete analysis of this dataset will be published elsewhere."

-> The two Appendix tables need to be updated to fit the journal nomenclature. If you want to keep them as part of the Appendix then please insert them in the Appendix pdf (and include them in the table of contents). Alternatively, you can keep them in the excel format by making them Expanded View tables (please also update the callouts in the main manuscript file if you do so). See our online author guidelines for more detail http://emboj.embopress.org/authorguide.

The two appendix tables are now part of the Appendix PDF and included in the table of contents

-> I am not sure if the browser screenshot in fig 3A is of sufficiently high resolution to display well in the final version. Is it possible for you to include a version with either higher resolution or a clearer colour scheme?

We have now provided new browser screenshot at good resolution and with a clear background.

-> The GFP-WB in figure 1D is extremely pixelated, I would suggest that you make the panel slightly smaller or use a higher resolution image for this

We have now used a higher resolution picture to make the new 1D figure.

Referee #2:

The manuscript by Deleris et al. has been substantially improved considering most of my previously raised points, and also much of the other reviewers' input. I admit that I am still a bit concerned about the high variability between experiments, but the authors have tried hard to reduce it and provide arguments why this is a difficult issue. I do not question the conclusions drawn on the data. I appreciate the much improved documentation of the methylation sites. I hope that the planned elimination of the solo-LTR will work in future and will strengthen the evidence for its role.

Referee #3:

This revised ms is an interesting and substantial piece of work. Authors have improved the data analysis and writing, and now place their findings more clearly in the context of other

published work. Authors have responded satisfactorily to my specific comments on the importance and generality of LTR locations to immunity gene expression control, and Hpa infection experiments. Generally, the data are carefully reflected on, and I like the notion that LTRs might contribute to promoter maturation for balanced cis-control of immunity/stress-related genes. Inclusion of Fig. 5F showing responsiveness of RPP4 gene in the various rpp4 LTR/W1-box transgenics helps reinforce the PAMP effects on expression of this gene.
A few minor points:

1. I would remove words like 'interestingly' from text.

The word “interestingly” has now been removed from text in most instances (pages 6,7,9,10 and 15)

2. I don't see Waese et al (2017) and Ordonnez et al (2013) in refs - end 1st paragr. Pg 10. Can authors carefully check through references.

This has been amended.

3. Figure S2 in Suppl. Materials is marked as Fig. S1. Can authors provide a short explanation to data in Appendix Tables S1 and S2.

This has now been amended in the legends.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Angélique Deleris
Journal Submitted to: EMBO journal
Manuscript Number: EMBOJ-2017-98482R1

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	na
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	na
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	na
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	na
For animal studies, include a statement about randomization even if no randomization was used.	na
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	na
4.b. For animal studies, include a statement about blinding even if no blinding was done	na
5. For every figure, are statistical tests justified as appropriate?	yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	yes; this was assessed with Graph Pad Prism software; if not, both T-test and non parametric tests were performed on the data sets, leading to the same statistical differences (detailed in Source Data)
Is there an estimate of variation within each group of data?	yes

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	yes if applicable, if not a Welch 's t-test was used

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	yes
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	na

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	na
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	na
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	na

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	na
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	na
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	na
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	na
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	na
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	na
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	na

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	na
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	na
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	na
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	na

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	na